

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

2100 2152 2430

520.41201X00

#3

Applicant(s): F. MACIEL

Serial No.: 10 / 084,474

Filed: FEBRUARY 28, 2002

Title: INFORMATION PROCESSING SYSTEM ACCESSED THROUGH
NETWORK AND CONTROL METHOD OF PACKET TRANSFER
LOAD

RECEIVED

APR 16 2002

Technology Center 2100

LETTER CLAIMING RIGHT OF PRIORITY

Assistant Commissioner for
Patents
Washington, D.C. 20231

MARCH 27, 2002

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s)
the right of priority based on:

Japanese Patent Application No. 2001-361913
Filed: NOVEMBER 28, 2001

A certified copy of said Japanese Patent Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP

Carl I. Brundidge
Registration No. 29,621

CIB/rp
Attachment



日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2001年11月28日

出願番号

Application Number

特願2001-361913

[ST.10/C]:

[JP2001-361913]

出願人

Applicant(s):

株式会社日立製作所

RECEIVED

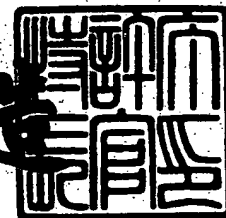
APR 16 2002

Technology Center 2100

2002年 3月 1日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2002-3011692

【書類名】 特許願

【整理番号】 H01004531A

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/00

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

 【氏名】 マシエル フレデリコ

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社 日立製作所

【代理人】

 【識別番号】 100075096

 【弁理士】

 【氏名又は名称】 作田 康夫

 【電話番号】 03-3212-1111

【手数料の表示】

 【予納台帳番号】 013088

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ネットワーク連携情報処理システムおよびその複数負荷分散機間のアクセス移動方法

【特許請求の範囲】

【請求項1】

外部ネットワークを介して伝達されるパケットの示すアクセスを実行するための情報処理装置群と、

前記外部ネットワークを介して伝達されるパケットを受信し、該パケットの情報のうち少なくとも受信者アドレスを前記情報処理装置群のうちのアクセスを実行すべき情報処理装置のアドレスに変換して通信パケットを転送し、もって通信を中継する並列動作可能な複数のネットワークドレス変換装置と、

前記ネットワークアドレス変換装置の各々に対応してアクセス対応表を保持し、該アクセス対応表の各エントリーには対応するネットワークアドレス変換装置が中継するアクセスの要求元と、該アクセスの要求先と、該アクセスの要求先に対応するネットワークドレス変換装置の外部インターフェースのアドレスとを記録する手段と、

ある情報処理装置に対するアクセスに関するパケットの中継の分担を第1のネットワークアドレス変換装置から第2のネットワークアドレス変換装置に移動すべきことを決定すると、該情報処理装置に対するアクセスの通信を前記第1のネットワークアドレス変換装置から前記第2のネットワークアドレス変換装置に移動するとともに、前記第2のネットワークアドレス変換装置に対し、該第2のネットワークアドレス変換装置が受信した前記情報処理装置に対するアクセスに関するパケットをメモリに保存するよう指示し、前記第1のネットワークアドレス変換装置に対応するアクセス対応表のエントリーの内、移動対象のアクセスに関するエントリーを前記第2のネットワークアドレス変換装置に対応するアクセス対応表に移動し、その後前記第2のネットワークアドレス変換装置に対し、前記保存したパケットの転送を前記第2のネットワークアドレス変換装置に対応するアクセス対応表にしたがって再開するように制御する制御機構とを備えたことを特徴とするネットワーク連携情報処理システム。

【請求項2】

前記制御機構は前記複数の情報処理装置のいずれが一つに設けられることを特徴とする請求項1記載のネットワーク連携情報処理システム。

【請求項3】

前記複数のネットワークアドレス変換装置は、前記外部ネットワークを介して伝達されるパケットの受信者アドレスおよび送信者アドレスを、アクセスを実行すべき情報処理装置のアドレスおよび自己の内部インターフェースのアドレスにそれぞれ変換して該パケットを転送し、

前記アクセス対応表には、中継するアクセスの要求先に対応するネットワークアドレス変換装置の内部アドレスが更に記録されることを特徴とする請求項1に記載のネットワーク連携情報処理システム。

【請求項4】

前記アクセス対応表は対応するネットワークアドレス変換装置の内部に各々保持され、かつ該アクセス対応表の各エントリーの内容はそのエントリーが示すアクセス対象の情報処理装置内のコピー保管機構に保存され、前記制御機構によるエントリーの移動の際には、前記コピー保管機構に保存されたエントリーの内容が移動先のネットワークアドレス変換装置の内部に保持されたアクセス対応表に複写されることを特徴とする請求項1に記載のネットワーク連携情報処理システム。

【請求項5】

外部ネットワークを介して伝達されるパケットの示すアクセスを実行するためのサーバ群と、

前記外部ネットワークを介して伝達されるクライアントからのパケットを受信し、該パケットの情報のうち少なくとも受信者アドレスをサーバ群のうちのアクセスを実行すべきサーバのアドレスに変換して通信パケットを転送し、もって通信を中継する並列動作可能な複数の負荷分散機と、

前機負荷分散機の各々に対応してアクセス対応表を保持し、該アクセス対応表の各エントリーには対応するネットワークアドレス変換装置が中継するアクセスの要求元と、該アクセスの要求先と、該アクセスの要求先に対応するネットワー

クドレス変換装置の外部インターフェースのアドレスとを記録する手段と、あるサーバに対するアクセスに関するパケットの中継の分担を第1の負荷分散機から第2の負荷分散機に移動すべきことを決定すると、該サーバに対するアクセスの通信を前記第1の負荷分散機から前記第2の負荷分散機に移動するとともに、前記第2の負荷分散機に対し、該第2の負荷分散機が受信した前記サーバに対するアクセスに関するパケットをメモリに保存するよう指示し、前記第1の負荷分散機に対応するアクセス対応表のエントリーの内、移動対象のアクセスに関するエントリーを前記第2の負荷分散機に対応するアクセス対応表に移動し、その後前記第2の負荷分散機に対し、前記保存したパケットの転送を前記第2の負荷分散機に対応するアクセス対応表にしたがって再開するように制御する制御機構とを備えたことを特徴とするネットワーク連携情報処理システム。

【請求項6】

外部ネットワークを介して伝達されるパケットを受信し、該パケットの示すアクセスを実行するための情報処理装置群の一つに該パケットを転送するネットワークアドレス変換装置の動作方法であって、

受信したパケットが前記情報処理装置の一つに対するアクセスを開始するために接続確立要求するパケットである時、該アクセスの要求元と、該アクセスの要求先と、該アクセスの要求先に対応する前記ネットワークアドレス変換装置の外部インターフェースのアドレスとの対応をアクセス対応表の一つのエントリーとして登録し、

受信したパケットのアクセスが前記アクセス対応表に登録されているとき、前記アクセス対応表にしたがってパケットの受信者アドレスを変換して、変換したパケットを転送し、

受信したパケットのアクセスが前記アクセス対応表に登録されていない時には転送を拒絶して該パケットの送信元にエラー返送し、

外部からアドレスを特定して指定があった時、該特定されたアドレスを受信者アドレスとするパケットを受信した場合は該受信者アドレスが前記アクセス対応表に登録されていなくとも送信元へのエラーの返送をせず、該パケットをメモリに保存することを特徴とするネットワークアドレス変換装置の動作方法。

【請求項 7】

前記アクセス対応表のエントリーの内容をそのエントリーに登録されたアクセス要求先に転送する過程をさらに有することを特徴とする請求項 7 記載のネットワークアドレス変換装置の動作方法。

【請求項 8】

外部ネットワークを介して伝達されるパケットの示すアクセスを実行するための情報処理装置群と、前記外部ネットワークを介して伝達されるパケットを受信し、該パケットの情報のうち少なくとも受信者アドレスを前記情報処理装置群のうちのアクセスを実行すべき情報処理装置のアドレスに変換して通信パケットを転送し、もって通信を中継する並列動作可能な複数のネットワークアドレス変換装置と、前記ネットワークアドレス変換装置の各々に対応してアクセス対応表を保持し該アクセス対応表の各エントリーには対応するネットワークアドレス変換装置が中継するアクセスの要求元と、該アクセスの要求先と、該アクセスの要求先に対応するネットワークアドレス変換装置の外部インターフェースのアドレスとを記録する手段とを有するシステムにて、前記複数のネットワークアドレス変換装置の間で通信負荷分散もしくはフェールオーバーのためにアクセスの中継の機能を受け渡すアクセス移動方法であって、

ある情報処理装置に対するアクセスに関するパケットの中継の分担を第 1 のネットワークアドレス変換装置から第 2 のネットワークアドレス変換装置に移動すべきことを決定すると、

前記第 2 のネットワークアドレス変換装置に対し、該第 2 のネットワークアドレス変換装置が前記情報処理装置に対するアクセスに関するパケットを受信した場合に、該パケットを転送せず、拒絶せず、メモリに保存するという移動モードへの移行を指示し、

前記情報処理装置に対するアクセスの通信を前記第 1 のネットワークアドレス変換装置から前記第 2 のネットワークアドレス変換装置に移動し、

前記第 1 のネットワークアドレス変換装置に対応するアクセス対応表のエントリーの内、移動対象のアクセスに関するエントリーを前記第 2 のネットワークアドレス変換装置に対応するアクセス対応表に移動し、

後前記第2のネットワークアドレス変換装置に対し、該第2のネットワークアドレス変換装置が前記情報処理装置に対するアクセスに関するパケットを受信した場合に、前記第2のネットワークアドレス変換装置に対応するアクセス対応表にしたがって該アクセスを転送するという通常動作モードへの復帰を指示し、前記保存したパケットについても前記アクセス対応表にしたがって転送させる、との手順を有するネットワークアドレス変換装置間のアクセス移動方法。

【請求項9】

前記アクセス対応表は、対応するネットワークアドレス変換装置の内部に各々保持され、かつ該アクセス対応表の各エントリーの内容はそのエントリーが示すアクセス対象の情報処理装置内のコピー保管機構に保存され、前記移動対象のアクセスに関するエントリーの移動の際には、前記コピー保管機構に保存されたエントリーの内容を移動先のネットワークアドレス変換装置の内部に保持されたアクセス対応表に複写することを特徴とする請求項8に記載のネットワークネットワークアドレス変換装置間のアクセス移動方法。

【請求項10】

外部ネットワークを介して伝達されるパケットの示すアクセスを実行するための情報処理装置群と、前記外部ネットワークを介して伝達されるパケットを受信し、該パケットの情報のうち少なくとも受信者アドレスを前記情報処理装置群のうちのアクセスを実行すべき情報処理装置のアドレスに変換して通信パケットを転送し、もって通信を中継する並列動作可能な複数のネットワークアドレス変換装置と、前記ネットワークアドレス変換装置の各々に対応してアクセス対応表を保持し該アクセス対応表の各エントリーには対応するネットワークアドレス変換装置が中継するアクセスの要求元と、該アクセスの要求先と、該アクセスの要求先に対応するネットワークアドレス変換装置の外部インターフェースのアドレスとを記録する手段とを有するシステムにて、前記複数のネットワークアドレス変換装置の間で通信負荷分散もしくはフェールオーバーのためにアクセスの中継の機能を受け渡すアクセス移動制御プログラムであって、

ある情報処理装置に対するアクセスに関するパケットの中継の分担を第1のネットワークアドレス変換装置から第2のネットワークアドレス変換装置に移動すべ

きことを決定すると、

前記第 2 のネットワークアドレス変換装置に対し、該第 2 のネットワークアドレス変換装置が前記情報処理装置に対するアクセスに関するパケットを受信した場合に、該パケットを転送せず、拒絶せず、メモリに保存するという移動モードへの移行を指示するステップ、

前記情報処理装置に対するアクセスの通信を前記第 1 のネットワークアドレス変換装置から前記第 2 のネットワークアドレス変換装置に移動するステップ、

前記第 1 のネットワークアドレス変換装置に対応するアクセス対応表のエントリーの内、移動対象のアクセスに関するエントリーを前記第 2 のネットワークアドレス変換装置に対応するアクセス対応表に移動するステップ、

後前記第 2 のネットワークアドレス変換装置に対し、該第 2 のネットワークアドレス変換装置が前記情報処理装置に対するアクセスに関するパケットを受信した場合に、前記第 2 のネットワークアドレス変換装置に対応するアクセス対応表にしたがって該アクセスを転送するという通常動作モードへの復帰を指示し、前記保存したパケットについても前記アクセス対応表にしたがって転送させるステップ、

を有するネットワークアドレス変換装置間のアクセス移動制御プログラム。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、複数の負荷分散機もしくはネットワークアドレス変換装置を備えたネットワーク連携情報処理システムに係り、特に複数のサーバへのアクセスの中継を分担する上記複数の負荷分散機もしくはネットワークアドレス変換装置の間で通信負荷分散もしくはフェールオーバを実現するためのアクセス中継の移動に関する。

【 0 0 0 2 】

【従来の技術】

現在インターネットでの通信量が急激に増加し、Web サイトは一台の Web サーバでアクセス数に対応できなく、複数の Web サーバを使用する。これらのサ

ーバにアクセスを分散するために様々な方法が使用されるが、近年負荷分散機という装置の利用が増えてきた。図1に負荷分散機の使用例を示す。クライアント1a～1cはインターネット2を介してサイトをアクセスする。このサイトでは負荷分散機3aはアクセスをサーバの代わりに受け、ネットワーク4を介してWebサーバプログラムを実行している複数のサーバ5a～5cの間で分散する。負荷分散機3はアクセス対応表9aを用い、クライアント1a～1cとサーバ5a～5cとの通信のパケットのネットワークアドレスをトランスペアレントに変換して負荷分散機能を実現する。アドレス変換の基本方式は“The IP Network Address Translator (NAT)”, Internet Engineering Task Force RFC1631 (今後参考文献1と呼ぶ)にて説明される。次にアドレス変換を説明する。本明細書ではネットワークインタフェースの識別子(10a～10c、31a、32a、51a～51c等)に「IP」を追加しIPアドレスを表現する。そして、上記のインタフェースは複数のIPアドレスがある場合、更に序数を追加しアドレスを特定する。

図1は一台の負荷分散機3aの利用を示す。しかし、アクセス数の増加のため、この負荷分散機3aがボトルネックになると考えられる。そして、負荷分散機3aが故障した場合クライアント1a～1cはサイト全体をアクセスできなくなるため、1台の負荷分散機3aの利用は可用性を低下する。このため、図2に示すとおり複数の負荷分散機3a、3bの並行利用が望ましい。

複数の負荷分散機の並行利用としてActive/StandbyとActive/Activeの2つの方式がある(「機能強化が進むWWWサーバ負荷分散装置」、日経オープンシステム1999年12月号、ISSN 0918-581X、128～131ページ参考、今後参考文献2と呼ぶ)。Active/Standby方式では1台の負荷分散機(例えば、3a)が動作し、残りの負荷分散機(例えば、3b)がスタンバイするため、複数台使用しても一台の性能を超えられない。一方、Active/Active方式の方には全ての負荷分散機が同時に動作するため効率が高い。

【0003】

【発明が解決しようとする課題】

しかしながら、従来のActive/Active方式は3つの問題がある。第一に、参考文献2にも指摘されているとおり負荷分散機間の通信を分散できない点である。クライアントは通常時、接続先の負荷分散機を固定して利用する。このため、動的に負荷分散機の間で通信負荷を分散できない。

第二の問題は、負荷分散機が故障して他の負荷分散機にアクセスをフェールオーバーするときには、故障した負荷分散機のアクセス情報がなくなるため、これらのアクセスが途中で中断される点にある。

第三の問題は、場合により負荷分散機の間で専用接続や、常時アクセス対応表9の他負荷分散機への複写という機能が使用されるが、負荷分散機の数が多くなるとこれらの機能はスケーラビリティを制限するという点である。

なお、これら3つの問題点は負荷分散機間（ネットワークアドレス変換装置間）で生じるだけでなく、通信を中継するネットワークアダプタもしくはゲートウェイ装置でも、複数のこの種の装置をActive/Active方式で並列にもちいる際に共通の問題である。

そこで本発明の一つの目的は、負荷分散機に代表されるネットワークアドレス変換装置もしくは通信を中継するネットワークアダプタもしくはゲートウェイ装置を複数個Active/Active方式で用いるシステムにて上記装置間の動的通信負荷分散を実現することにある。

本発明の他の目的は、上記複数の装置間でアクセスの中断を伴うことなく通信の移動を実現すること、別の表現では、フェールオーバーを可能にすることにある。

本発明のさらに別の目的は、Active/Active方式で用いる上記装置の個数を容易に増減できるようにし、つまりスケーラビリティを増大させることにある。

【0004】

【課題を解決するための手段】

本発明の代表的態様にしたがうネットワーク連携情報処理システムは、ネットワークを介して伝達されるパケットの示すアクセスを実行するための複数の情報処理

装置と、パケットを受信してアクセスを実行すべき情報処理装置にパケットを転送する複数のネットワークアドレス変換装置（NAT装置）を含み、第1のネットワークアドレス変換装置から第2のネットワークアドレス変換装置へ、ある情報処理装置に対するアクセス中継の分担を移動する場合に、パケットの変換およびアクセスの応答の返送に用いるため個々のネットワークアドレス変換装置にて保有するアクセス対象表の移動対象アクセスに関するエントリーを第1のネットワークアドレス変換装置から第2のネットワークアドレス変換装置に移動し、次にこの情報処理装置へのアクセスの通信を第1のネットワークアドレス変換装置から第2のネットワークアドレス変換装置に移動する。この二つの移動動作の間にパケットがネットワークから到達してしまう事態に対処するため、アクセス中継の移動先である第2のネットワークアドレス変換装置の動作モードを、上記情報処理装置へのアクセスを示すパケットが受信されると、アクセス対応表に未だ登録がないからといって拒絶することを行わず、そのパケットをメモリに保存する移動モードに予め変更しておく。通信の移動が完了した後に、第2のネットワークアドレス変換装置の動作モードを通常動作モード、すなわち自身のアクセス対応表にしたがってアクセスを実行すべき情報処理装置に受信したパケットを転送し、受信したパケットの示すアクセスがアクセス対応表に登録されていない時にはこのパケットを拒絶し、送信元にエラーを返送するという動作モードに復帰させる。

ネットワークアドレス変換装置間のフェールオーバーを実現する特徴的構成は、上記アクセス対応表の内容を第3者に保存しておく点にある。あるネットワークアドレス変換装置の故障検出によりそのネットワークアドレス変換装置から別のネットワークアドレス変換装置にアクセス中継の機能を移動すべきと決定した時、アクセス対応表間の移動対象アクセスに関するエントリーの移動は、その第3者から移動先のネットワークアドレス変換装置のアクセス対象表へ複写することにより実行する。代表的には各アクセス対応表の各エントリーの内容は、そのエントリーのアクセスの対象である情報処理装置に保存しておく。

上記アクセス中継の移動の手順を制御する制御機構は、複数のネットワークアドレス変換装置に関わるので、通常はネットワークアドレス変換装置の外部に設け

る。複数の情報処理装置のうちの一つにこの制御機構を設けてもよい。

一方、個々のネットワークアドレス変換装置、もしくは負荷分散機として特徴的な点は、アドレスを変換してパケットを転送する機能、アクセス対応表に登録されていないアクセス受信したとき、このアクセスを拒絶してエラーを転送する機能を外部からの指定によりオフする点にある。

【 0 0 0 5 】

【 発明の実施の形態 】

<< 発明の実施の形態 1 >>

図 3 は実施の形態 1 のネットワーク連携情報処理システムの構成を示す。複数の Web サーバ 5 a、5 b、5 c が内部ネットワーク 4 により複数の負荷分散機 3 a、3 b に接続され、これによりインターネットでのあるサイトを実現するネットワーク連携処理システムが構成される。

サイトには通常、負荷分散機 3 a、3 b、ネットワーク 4、サーバ 5 a ~ 5 c の他に多くの装置があると考えられるが、図では本発明に関わる装置のみを示す。そして、負荷分散機は Web アクセスに限られておらず、FTP や電子メールなどの他のインターネットサービスにも利用でき、サーバ 5 a ~ 5 c も Web 以外のサービスも実行することがある。

本実施の態様に特徴的な負荷分散機間のアドレス中継機能の移動を説明する前に、図 4 と図 5 を参照して負荷分散機のアドレス変換を説明する。図 4 (a) にクライアント 1 a と負荷分散機 3 a の間のパケット、図 4 (b) に負荷分散機 3 a とサーバ 5 a の間の同じパケットを示す。各パケットは送信者の IP アドレス 8 0 0、8 0 5 とポート番号 8 0 1、8 0 6、受信者の IP アドレス 8 0 2、8 0 7 とポート番号 8 0 3、8 0 8 とデータ等の他の情報 8 0 4、8 0 9 を持つ。負荷分散機 3 a は図 4 (a) のパケットをクライアント 1 a から受信した場合、送信者と受信者アドレス 8 0 0 ~ 8 0 3 からアクセスを特定し、図 4 (b) に示すとおり受信者 IP アドレス 8 0 7 をこのアクセスを処理しているサーバ (この例では 5 a) の IP アドレスに変更し、パケットをネットワーク 4 に送信する。サーバ 5 a は受信者アドレス 8 0 7 のためこのパケットを受信する。逆方向の通信についてまず、サーバ 5 a ~ 5 c はインターネット 2 へのルータとして負荷分散

機 3 a を使用するため、クライアント 1 a への返信パケットを負荷分散機 3 a に送信する。負荷分散機 3 a は返信パケットをクライアント 1 a に送信する前に逆変換を行い、図 4 (a) のアドレス 802 とする。

負荷分散機 3 a が上記の変換を行うために図 5 に示すアクセス対応表 9 a を使用する。アクセス対応表 9 a はそれぞれのアクセスの情報を 1 行に保存する。この情報はクライアント 1 a の IP アドレス 901 とポート番号 902、負荷分散機の IP アドレス 903、そしてこのアクセスを処理しているサーバ 5 a ~ 5 c の IP アドレス 904 とポート番号 905、そして TCP フロー制御情報 906 ~ 908 である（参考文献 1 は TCP フロー制御情報の詳細を説明するため、本明細書では説明を省く）。負荷分散機 3 a がクライアント 1 a からパケットを受信したとき（図 4 (a) ）、800 ~ 803 の情報を 901、902、903、905 との対応でアクセスを特定し、そしてこのアクセスを 904 の IP アドレスに変換し図 4 (b) のパケットを構成する。負荷分散機は逆方向の通信の変換も同様に行う。負荷分散機 3 a がアクセス対応表 9 a にないアクセスのパケットを受信したとき、このパケットがアクセスの始まりである接続確立要求の場合にアクセス対応表 9 にこのアクセスを追加し、そうでなければこのパケットの送信者（クライアント 1 a ~ 1 c かサーバ 5 a ~ 5 c）にエラーを返す。そして負荷分散機 3 a はアクセス終了時にこのアクセスの行をアクセス対応表 9 a から削除する。

負荷分散機間通信負荷分散およびフェールオーバーは 2 つのステップ、すなわち割り当ての計算と通信の移動とで実現できる。割り当ての計算のステップでは、どの通信をどの負荷分散機に割り当てると通信負荷が分散するかを計算する。また、故障した負荷分散機に通信を割り当てないことでフェールオーバーを実現する。一方、通信の移動のステップでは、通信を実際に移動し、この割り当てを実現する。

割り当ての計算の例は「“Dynamic Gateways: A Novel Approach to Improve Networking Performance and Availability on Parallel Servers,” Proceedings of the HPCN

‘98, pp 678-687, Springer-Verlag, 1998, ISSN 0302-9743」(今後参考文献3と呼ぶ)、特開平10-224400そして米国特許6112248に詳しく紹介されている。

通信の移動については、まず従来技術の問題点を説明する。アクセスを負荷分散機3aから3bに移動するために、このアクセスの通信と、アクセス対応表9aでのこのアクセス情報を負荷分散機3bに移動する必要がある。しかし、この2つの間で「ひよこが先か卵が先か」という問題が生じる。通信を移動した後にアクセス情報を移動した場合、この2つの移動の間に負荷分散機3bがアクセス対応表9bに登録していないアクセスを受信するため、送信元にエラーを返す。逆に、アクセス情報を移動した後に通信を移動した場合、この2つの移動の間に負荷分散機3aでのアクセス情報が変化されるため、通信を移動したときに負荷分散機3bでのアクセス情報が古い(もう存在しないアクセスがまだ登録され、新しいアクセスがまだ登録されていない)、等の問題が生じる。

本実施の態様における通信の移動(負荷分散機間のアクセス中継機能の移動)を次に説明する。アクセス移動前の負荷分散機3aの内部のアクセス対応表9aの内容は先に示したを図5の通りとする。また、同じくアクセスを移動前の負荷分散機3bのアクセス対応表9bの内容を図6に示す通りとする。クライアント1aとサーバ5aの間のアクセスの中継を負荷分散機3aから3bに移動することを例にして説明する。

ここで本実施の形態でのアクセス中継の移動の概要をまず述べると、移動の単位は、各負荷分散機のインターネット2に対するインターフェース(外部インターフェース)のIPアドレス単位である。したがって、外部インターフェース31aのIPアドレス31a-IP-1もしくは31a-IP-2、または外部インターフェース31bのIPアドレス31b-IP-1などが移動の単位とする。

さらに本実施の態様では、複数のサーバの各々は、それら外部インターフェースのIPアドレスにそれぞれ所属している。図の例では、サーバ5a、5b、5cをそれぞれIPアドレス31a-IP-1、31a-IP-2、31b-IP-1に所属している。この場合クライアントからインターネット2を介して伝達されるパケットの受信者IPアドレスは図2のとおり外部インターフェースのIP

アドレスである。負荷分散機では、その受信者IPアドレスをそのIPアドレスに対応する（IPアドレスが所属する）サーバのIPアドレスに変換して、パケットをサーバに転送する（図4の（b））。

したがって、外部インターフェースのIPアドレス（例えば31a-IP-1）の割り付けを負荷分散機3aから負荷分散機3bに移動することで、このIPアドレスが所属するサーバに対するアクセスの中継を負荷分散機3bに移動することができる。なお、この方法では、負荷分散機の数より多くのIPアドレス数を使用する必要がある。このIPアドレスの間の通信負荷分散についてはRound-robin DNS (Eric Dean Katz, Michelle Butler, Robert McGrath, "A Scalable HTTP Server: The NCSA Prototype," Proceedings of the First International Conference on the World-Wide Web, 1994参照) に記載された技術を使用できる。

さて本実施の態様では、図3のサーバ5cに設けた制御機構52が、各負荷分散機に割り付けたIPアドレスの管理、負荷分散機間通信負荷分散のための各負荷分散機の通信量の情報収集、割当の計算、および上記したIPアドレスの移動によるアクセス中継移動の指示を行う機能を備えている本実施例では、制御機構52による割り当ての計算の結果、負荷分散機の間でのIPアドレスの割り当てが分かる。そして、割り当ての変更で移動が分かる。さて本実施例に特徴的なのは、このIPアドレスの移動による通信の移動をどのように実現するかという点にある。ここでも、例としてIPアドレス31a-IP-1を負荷分散機3aから3bに移動するのを例にとって説明する。すなわち、図5と図6では、当初IPアドレス31a-IP-1と31a-IP-2が負荷分散機3aに属しており、IPアドレス31b-IP-1が負荷分散機3bに属している。制御機構52の、割り当ての計算の結果が31a-IP-2を負荷分散機3a、31a-IP-1と31b-IP-1を負荷分散機3bに割り当てるべきとの結果であった。その結果、図3と図6で示すアクセス移動前の状態と比べて、IPアドレス31a-IP-1を負荷分散機3bに移動せねばならない。

IPアドレスの移動を実現する方法は次の4つのステップからなる。

第一ステップでは、制御機構52は負荷分散機3bに31a-IP-1の移動を知らせる。この知らせを受けると、負荷分散機3bはこのアドレスに関して通常動作モードとは異なる移動モードにと設定する。切り換える。この移動モードは、本発明に特有のモードであり、指定されたIPアドレスのパケットを受信した場合、アクセス対応表9a、9bに新アクセスとして登録せず、そしてエラーを返さず、パケットをメモリに保存するモードである。なお、実際には、次に述べる第二ステップの通信の移動の後で始めて、指定されたIPアドレスのパケットが移動させるべき負荷分散機3bに到達する場合が生じる。また上記の通常動作モードとは、そのモードであると指定されたIPアドレスのパケットを受信した場合に図4にて説明したIPアドレスの変換、パケットの転送を実行し、またもしアクセスの開始のために接続確定要求を示すパケットを受信した場合は、その後アクセスを継続するためのパケットの転送もしくはサーバからのアクセス応答の返送に備えてアクセス対応表にそのアクセスを登録する動作モードである。もし、ある負荷分散機に移動モードとしての指定も、通常動作モードとしての指定もないIPアドレスを受信者アドレスとするパケットが届いた場合、送信元にエラーを返す。これらの移動モード、通常動作モードのモード設定を実現するために、各負荷分散機3a、3bには処理モード表7a、7bを装備する。図7の(a)と(b)にそれぞれに移動前の処理モード表7a、7bの内容を、図8の(a)に31a-IP-1の移動を知らせた後の処理モード表7bの内容を示す。このように、各負荷分散機の処理モード表7a、7bには、IPアドレスに関する負荷分散機の処理モードとして通常動作モード、移動モードが設定されている期間中に、そのIPアドレス70ごとに処理モード71が保存される。なお、この第一のステップが終了しても、負荷分散機3aの処理モード表7aの上ではIPアドレス31a-IP-1は通常動作モードと設定されたままであり、IPアドレス31a-IP-1を受信者アドレスとする新規パケットは通常どうり負荷分散機3aを介して目的とするサーバ5aに転送される。

第二ステップでは、制御機構52の制御により31a-IP-1の通信を負荷分散機3aから3bに移動する。これは、負荷分散機のインターフェースへのIP

アドレス31a-IP-1の設定を解除し、負荷分散機3bのインターフェースに設定することにより実行する。つまりIPアドレス31a-IP-1に属しているサーバ5aのインターネット2へのルータを負荷分散機3bに変更する。この移動に参考文献3が説明するいくつかの実現方法(Proxy ARP、OSPF、サーバのルート変更)、そしてVRRPプロトコル(“Virtual Router Redundancy Protocol”, Internet Engineering Task ForceのRFC2338)を使用できる。第二ステップが完了すると、IPアドレス31a-IP-1を受信者アドレスとする新規パケットは負荷分散機3aではなく負荷分散機3bに届くようになる。この31a-IP-1を受信者アドレスとする新規パケットは、先の第一ステップでの移動モードの設定により、負荷分散機3bにてメモリに保存される。

第三ステップでは、制御機構52の制御により、負荷分散機3aのアクセス対応表9aの、負荷分散機IPアドレス903が31a-IP-1の値を持つ全ての行を、負荷分散機3bのアクセス対応表9bに移動する。すなわち負荷分散機3aのアクセス対応表9aの、負荷分散機IPアドレス903が31a-IP-1の値の行を負荷分散機3bのアクセス対応表9bにコピーする。図9にコピー後のアクセス対応表9bの内容を示す。その後、アクセス対応表9aからこれらの行を削除する。図10に移動後のアクセス対応表9aの内容を示す。

第四ステップでは、制御機構は31a-IP-1の移動終了の知らせを負荷分散機3a、3bに送信する。この知らせを受けると、負荷分散機3aは処理モード表7aから31a-IP-1を削除し、このアドレスに関するパケットを処理しなくなる。また、負荷分散機3bは、上記知らせを受けると、このアドレスに関する処理を移動モードから通常動作モードに切り換える。図8(b)、(c)のそれぞれに移動後の処理モード表7a、7bの内容を示す。そして、負荷分散機3bはメモリに保存していたパケットをアクセス対応表9bに従って処理する。すなわち、メモリから読み出したパケットの受信者アドレス31a-IP-1をキーにして、このパケットを処理すべきサーバのIPアドレス51a-IPを対応表9bで読み出し、パケットの受信者IPアドレスを51a-IPに変換する。

。変換したパケットはサーバ5aに送信される。

以上の手順の採用により、必要に応じてある負荷分散機のクライアント・サーバ間のアクセスの中継の機能を別の負荷分散機へ移動することができ、つまり複数の負荷分散機間での動的通信負荷分散が可能となる。

【0006】

<<発明の実施の形態2>>

参考文献1はアドレス変換の基本を説明するが、具体的に複数の変形のアドレス変換方法が利用される。本実施形態ではもう一つのアドレス変換方法における負荷分散機間通信負荷分散について述べる。

図11と図12に本実施形態のアドレス変換方式を示す。図11に示すとおり、本実施形態における負荷分散機（図5の3a、3bなど）のアドレス変換方法は受信者IPアドレス812、817だけでなく、送信者のIPアドレス810、815とポート番号811、816も変更する。このため、図3に示す実施形態1のアクセス対応表9aに比べて、本実施形態のアクセス対応表9a（図12）では負荷分散機3aの内部インタフェース32aのIPアドレス913とポート番号914をも追加して格納する点異なる。

負荷分散機からサーバへ送付するパケットの送信者アドレス815、816は負荷分散機3aもしくは3bのアドレスになるため、サーバ5a～5cにはクライアント1a～1cでなく負荷分散機3aもしくは3bがアクセスを送信しているように見える。このため、負荷分散機の外部インタフェース31a、31bに設定されたIPアドレスにサーバを所属させなくても良く、サーバの各々は負荷分散機の内部インターフェースのアドレスに属する。その内部アドレスの各々がそれぞれ設定された外部インターフェースのアドレスに属する。したがって、本実施形態のアドレス変換方法を採用すると、複数の負荷分散機3a、3bが一台のサーバにアクセスを送信することも可能となる。

さて本実施形態では、外部インタフェース31a、31bのIPアドレスに内部インタフェース32a、32bのIPアドレスを所属させて負荷分散機のアクセス対応表に格納しておく。したがって、先に示した実施形態1で第一ステップから第四ステップにて説明した他の負荷分散機へのアクセスの移動に際しては、外

部インタフェース 31a、31b の IP アドレスとこれに属する内部インタフェース 32a、32b の IP アドレスを一緒に移動する。これ以外、本実施形態は実施形態 1 と同じである。

【0007】

<<発明の実施の形態 3>>

以上の実施形態では負荷分散機間通信負荷分散方法を説明した。本実施形態では負荷分散機間フェールオーバー方法を説明する。

フェールオーバーの場合には通信負荷分散と同じく、アクセスをある負荷分散機から他の負荷分散機に移動する（以前の実施形態と同じく、ここでは負荷分散機 3a から負荷分散機 3b の移動を説明する）。しかし、フェールオーバーを必要とする障害が発生すると、多くの場合、その障害が発生した負荷分散機からアクセス対応表を読み出すことが不可能となる。このため、本実施形態では図 13 に示すとおり、各負荷分散機は通常の稼動状態時に、アクセス対応表の一部か全てを、アクセス対象のサーバのコピー保管機構 53 に送信する。各サーバのコピー保管機構 53 は、受信したアクセス対応表のコピー 54 を保存する。なお、図 13 では上記コピー保管機構 54 をサーバ 5a の内部にのみ図示しているが、アクセス対象となり得るサーバの全てに同様のコピー保管機構を設ける。

フェールオーバー時に、基本的には実施形態 1 で説明した 4 つのステップのアクセス移動の動作に従う。但し、第三ステップに示した、アクセス対応表同士のデータの移動が実施できる保証がないので、これに替えてアクセス対象のサーバに保存されたアクセス対応表のコピーの情報を移動先の負荷分散機のアクセス対応表に設定する。

フェールオーバーを必要とする障害が負荷分散機 3a に発生し、その結果、先の実施形態と同じく負荷分散機 3a から負荷分散機 3b へアクセスを移動する場合の 4 つのステップをあらためて記載する。

第一ステップでは、制御機構 52 からの通知により負荷分散機 3b は処理モード表 7b に IP アドレス 31a - IP-1 のアクセスが移動モードであることを登録する。

第二ステップでは IP アドレス 31a - IP-1 の設定を負荷分散機 5a から 5

bに移動し、IPアドレス31a-IP-1に関する通信を負荷分散機3aから3bに切り換える。

第三ステップでは、アクセス対応表9aのコピー54から移動すべきアクセスの情報を読み出し、負荷分散機3bのアクセス対応表9bに設定する。

第四ステップでは、制御機構52からアクセス移動終了の知らせを負荷分散機3a、3bに送信する。負荷分散機3aは処理モード表7aから31a-IP-1を

削除する。負荷分散機3bは処理モード表7bを書き換えてIPアドレス31a-IP-1の処理モードを移動モードから通常処理モードに切り換える。これにより負荷分散機7bはメモリに保存していたパケットを読み出し、アクセス対応表9bの情報にしたがってアドレスを変換して送出する。

以上のフェールオーバーのためのアクセス移動の処理と制御を制御機構52が実行するのは先の実施態様と同じである。本実施態様にて、アクセス対象サーバにアクセス対応表のコピー54として通常の稼働時から保存すべき内容について以下に述べる。まず、先の実施の態様1のように、負荷分散機でクライアントから送信されたパケットの受信者IPアドレスのみをアクセス対象サーバのIPアドレスに変換して転送するアドレス変換方法を採用するシステムの場合（図2を参照）には、図3のアクセス対応表のうちクライアントのアドレス901および902、サーバ側のポート番号905、そしてTCPフロー制御906、907、908をコピー54として保存する。サーバは一つの負荷分散機外部IPアドレス903に属しており、そしてサーバIPアドレス904はサーバ自身に属している。制御機構52はこれらの情報を持つためこれらIPアドレス903、904のコピーは不要である。

一方、クライアントから送信されたパケットの送信者IPアドレスと受信者IPアドレスと負荷分散機で変換し、負荷分散機を送信元として転送する実施の形態2で述べたアドレス変換方法を採用するシステムではアクセス対象サーバに保存するアクセス対象表の項目は、図12の各項目のうちサーバのIPアドレス915以外の全項目である。

【0008】

<<変形例>>

本発明はすでに記載した実施の形態あるいはその変形例に限定されるのではなく、以下に例示する変形例あるいは他の変形例によっても実現可能であることは言うまでもない。また、上記複数の実施の形態あるいはその変形例として記載の技術あるいは以下の変形例の組み合わせによっても実現できる。

(変形例1) 実施例1で説明したアドレス変換方法(図42)を使用し、そしてアクセス対応表9a、9bのTCPフロー制御情報906、907、908を必要としない通信プロトコル(HTTPなど)を使用した場合、フェールオーバーを実現するためにアクセス対応表9aの内容をアクセス対応表のコピー54にコピーすることは不要である。サーバ5aのオペレーティングシステムおよびまたはアダプタ51aに、図14に示すTCP/IPの接続を列挙するTCP/IP接続表が存在する。フェールオーバー時、第三ステップでは接続表の内容をアクセス対応表9bにコピーし、そしてサーバ5aが属しているIPアドレスを負荷分散機外部IPアドレスの欄903に書き込み、そしてdeltaの値907をゼロにする。

(変形例2) フェールオーバー時に、実施形態3で説明した4つのステップの代わりに、第三ステップ(アクセス対応表のコピー54か図14のTCP/IP接続表をアクセス対応表9bに設定すること)と第二ステップ(通信の移動)をこの順番で実行することも可能である。この場合、負荷分散機3a、3bを移動モードにする必要はない。

(変形例3) 負荷分散機3a、3bは移動モードでは、指定されたIPアドレスの packets を受信した場合、この packets をメモリに保存せず無視しても良い。この場合クライアント1a~1cとサーバ5a~5cが packets の損失を検出し再送する。

(変形例4) 本発明では負荷分散機3a、3bは負荷分散機に限られておらず、他のNAT装置にも適用できる。

(変形例5) 本発明はNATだけでなく、ネットワークアダプタにも利用できる。近年、InfiniBand等のネットワークが開発され、アダプタ等の装置は従来のサーバ内蔵、共有しない形態に制限しておらず、外付けおよびまたは共

有アダプタが可能になった (InfiniBand Trade Association, "InfiniBand Architecture Specification Volume 1 参照)。この例として、特開平10-69471は並列計算機やクラスタに接続するための共有ネットワークアダプタを説明する。特開平10-69471の図3と図4は外部ネットワークアドレス (や接続識別子) と内部バッファのアドレスを変換する表を示す。これらの表はアクセス対応表9と同様な位置付けである。例えば、本発明の図5や図13のネットワーク4が特開平10-69471のネットワークまたはInfiniBand、負荷分散機3a、3bが負荷分散機ではなく共有アダプタの場合、本発明は共有アダプタの間の通信負荷分散およびまたはフェールオーバーに利用できる。

(変形例6) 変形例5はパケットを転送するアダプタに関するが、本発明はプロトコルを処理するアダプタにも利用できる。近年、TCP/IP処理を行うアダプタが開発されてきた (Lucent Technologies, "Integrating the LAN, WAN & SAN for Optimized Network Performance," e-Commerce Infrastructure Technologies Conference and Tradeshow, Monterey, USA, February 2001 参照)。これらのアダプタは図14のTCP/IP接続表を内蔵する。本発明はこの場合のTCP/IP接続表のアダプタ間の移動にも使用できるため、アダプタ間の通信負荷分散およびまたはフェールオーバーに使用できる。

(変形例7) 本発明はTCP/IPだけでなく、他のプロトコルにも使用できる。そして本発明は、クライアント1a~1cと負荷分散機3a、3bの間で使用する通信プロトコルが、負荷分散機3a、3bとサーバ5a~5cの間で使用する通信プロトコルと同じであることに制限されておらず、それぞれのネットワークに異なったプロトコルを使用することが考えられる。この例として、「高速ソケット」という通信方式が挙げられる。高速ソケットの例として、公開特許公報特開平11-328134、Berkeley大学の方式 (S. H. Rodrigues, T. E. Anderson, D. E. Culler, "H

igh-Performance Local Area Communication With Fast Socket," Proceedings of the USENIX '97, 1997, pp. 257-274参照)、Shahらによる方式(H. V. Shah, C. Pu, R. S. Madukkarumukumana, "High Performance Sockets and RPC over Virtual Interface (VI) Architecture", Proceedings of CANPC' 99, 1999参照)が挙げられる。高速ソケットはアプリケーションの通信関連関数呼び出しをInfiniBand等のネットワークの高速通信機能にマッピングすることで高速通信を実現する。このため、IPでなく独自プロトコルを使用する。負荷分散機3a、3bがクライアント1a~1cとIPプロトコルで通信し、そしてサーバ5a~5cと高速ソケットで通信している場合、クライアント1a~1cのIPアドレスと高速ソケットで使用するアドレスを変換するために、アクセス対応表9a、9bと同様な表を使用する。この場合、アクセス対応表のサーバ側のアドレス(904、905、915、916)が高速ソケットで使用するアドレスを持つ。本発明は本変形例の場合の通信負荷分散およびまたはフェールオーバーを実現する。

(変形例8) 変形例7には、装置3a、3bが負荷分散機であることを述べた。しかし、これらの装置がインターネット2上の通常IP(や他のプロトコル)の通信と、ネットワーク4上の高速ソケット通信を変換し、負荷分散機能を含まないNAT装置およびまたはゲートウェイ装置の場合にも本発明を使用できる。

(変形例9) 図3と図13では制御機構52がサーバ5cで実行しているが、本発明はこの構成に制限されず、他のサーバ5a、5b、負荷分散機3a、3bや、図に示さない別の装置で実行することが考えられる。そして、制御機構52をハードウェアとしてもソフトウェアとしても実現することが考えられる。

(変形例10) 実施形態3ではコピー保管機構53とアクセス対応表のコピー54はサーバ5aにあるが、変形例9と同じくこれらは他の装置に存在することも可能である。

(変形例11) 本発明は参考文献3の通信負荷分散アルゴリズムに制限しておら

ず他の通信負荷分散アルゴリズムや発見的方法を利用できる。

(変形例 1 2) 図 5 と図 1 2 に示したアクセス対応表は、これらの図で示した欄の情報に制限しておらず、負荷分散機 3 a、3 b の機能により他の欄が存在する。図 1 4 も同じく、示した欄の情報に制限しておらずオペレーティングシステムやアダプタの機能により他の欄が存在する。

(変形例 1 3) 負荷分散機間通信負荷分散を行うとき、実施形態 1 に記述された第三ステップでは、アクセス対応表 9 b に設定する情報をアクセス対応表 9 a でなく、アクセス対応表のコピー 5 4 から読み出すことも可能である。

(変形例 1 4) 負荷分散機間フェールオーバーを行うとき、アクセス対応表 9 a の情報を読み出すことが可能な場合、アクセス対応表のコピー 5 3 でなくアクセス対応表 9 a の情報を読み出すアクセス対応表 9 b に設定することも可能である。

【 0 0 0 9 】

なお、本発明を実施するためのプログラムは、それ単独であるいは他のプログラムと組み合わせて、ディスク記憶装置等のプログラム記憶媒体に記憶された販売することができる。また、本発明を実施するためのプログラムは、すでに使用されている通信を行うプログラムに追加される形式のプログラムでもよく、あるいはその通信用のプログラムの一部を置換する形式のプログラムでも良い。

【 0 0 1 0 】

【発明の効果】

本発明はまず、負荷分散機間動的通信負荷分散を実現し、スケーラビリティ向上、自動チューニングによる効率向上、コスト削減の効果がある。そして、本発明は負荷分散機間フェールオーバーを実現し、システム全体の可用性を向上する。本発明ではあるクライアントの通信を他の負荷分散機に移動でき、クライアントは通常時、接続先の負荷分散機を固定して利用する問題を解決する。そして本発明の方法は通信負荷分散、フェールオーバー時以外では負荷分散機間通信が不要であり、負荷分散機の間専用接続も不要である。これらのため、多数の負荷分散機を並行に使用することができ、システム全体のスケーラビリティを向上する。そして、フェールオーバー時にアクセス中断がなく、アクセスデータがなくな

らなく電子取引等のサイトに適している。

【図面の簡単な説明】

【図1】

一台の負荷分散機を使用するサイトの構成を示すブロック図。

【図2】

複数台の負荷分散機を使用するサイト（従来方式）の構成を示すブロック図。

【図3】

実施態様1のサイトの構成を示すブロック図。

【図4】

実施態様1のアドレス変換法を示すデータ構成図。

【図5】

負荷分散機9aのアクセス移動前のアクセス対応表。

【図6】

負荷分散機9bのアクセス移動前のアクセス対応表。

【図7】

処理モード表（アクセス移動前）。

【図8】

処理モード表（アクセス移動の途中と後）。

【図9】

負荷分散機9bのアクセス移動後のアクセス対応表。

【図10】

負荷分散機9aのアクセス移動後のアクセス対応表。

【図11】

実施態様2のアドレス変換法を示すデータ構成図。

【図12】

図11のアドレス変換を行うためのアクセス対応表。

【図13】

フェールオーバーを実現する実施態様3のアクセス対応表コピー動作を示すブロック図。

【図14】

オペレーティングシステムのTCP/IP接続表。

【符号の説明】

1a～1c：クライアント

3a, 3b：負荷分散機

5a～5c：サーバ

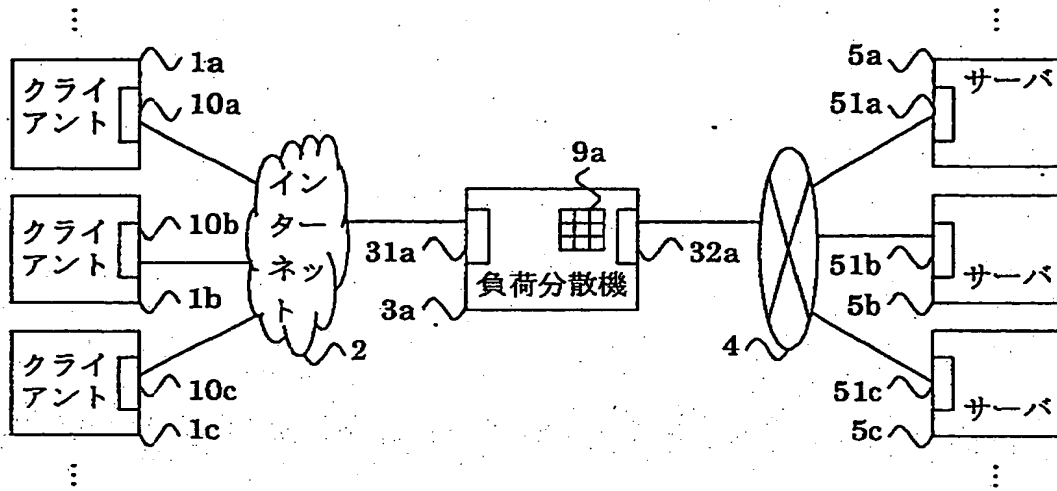
7a, 7b：処理モード表

9a, 9b：アクセス対応表。

【書類名】 図面

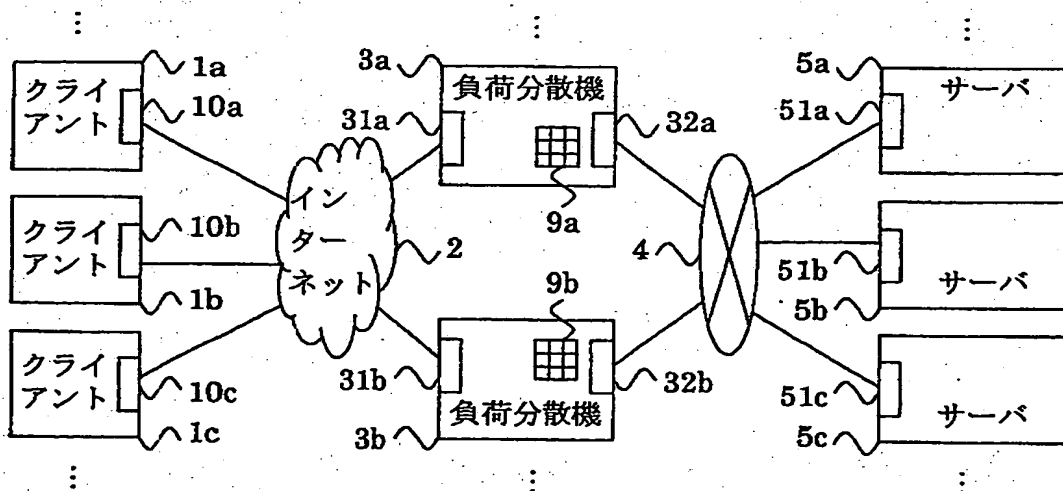
【図 1】

図 1



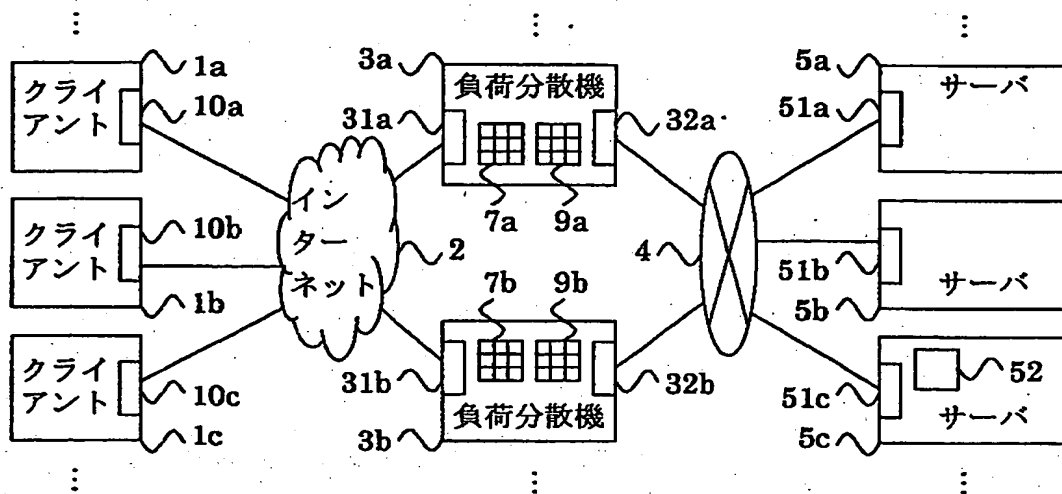
【図 2】

図 2



【図 3】

図 3



【図 4】

図 4

| 送信者アドレス (クライアント) | | 受信者アドレス (負荷分散機) | | 他 |
|---------------------|------|--------------------|------|-----|
| IP | port | IP | port | |
| 10a-IP | 1029 | 31a-IP-1 | 80 | ... |

800 801 802 803 804

(a) クライアントと負荷分散機の間

| 送信者アドレス (クライアント) | | 受信者アドレス (サーバ) | | 他 |
|---------------------|------|------------------|------|-----|
| IP | port | IP | port | |
| 10a-IP | 1029 | 51a-IP | 80 | ... |

805 806 807 808 809

(b) 負荷分散機と対象ホストの間

【図 5】

図 5

| クライアント | | 負荷分散機 (外部) | | サーバ | | TCP フロー制御 | | |
|--------|------|---------------|--|--------|------|-----------|-------|-----------|
| IP | port | IP | | IP | port | seq. | delta | timestamp |
| 10a-IP | 1029 | 31a-IP-1 | | 51a-IP | 80 | 320 | 0 | — |
| 10b-IP | 1028 | 31a-IP-2 | | 51b-IP | 21 | 29 | 2 | — |
| ... | ... | ... | | ... | ... | ... | ... | ... |

901 902 903 904 905 906 907 908

【図 6】

図 6

| クライアント | | 負荷分散機 (外部) | サーバ | | TCP フロー制御 | | |
|--------|------|---------------|--------|------|-----------|-------|-----------|
| IP | port | IP | IP | port | seq. | delta | timestamp |
| 10c-IP | 1050 | 31b-IP-1 | 51c-IP | 80 | 89 | 0 | — |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

901 902 903 904 905 906 907 908

【図 7】

図 7

| IP | モード | IP | モード |
|----------|------|----------|------|
| 31a-IP-1 | 通常動作 | 31b-IP-1 | 通常動作 |
| 31a-IP-2 | 通常動作 | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |

70 71 70 71

(a) (b)

【図 8】

図 8

| IP | モード | IP | モード | IP | モード |
|----------|------|----------|------|----------|------|
| 31b-IP-1 | 通常動作 | 31a-IP-2 | 通常動作 | 31b-IP-1 | 通常動作 |
| 31a-IP-1 | 移動 | ⋮ | ⋮ | 31a-IP-1 | 通常動作 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

70 71 70 71 70 71

(a) (b) (c)

【図 9】

図 9

| クライアント | | 負荷分散機 (外部) | サーバ | | TCP フロー制御 | | |
|--------|------|---------------|--------|------|-----------|-------|-----------|
| IP | port | IP | IP | port | seq. | delta | timestamp |
| 10c-IP | 1050 | 31b-IP-1 | 51c-IP | 80 | 89 | 0 | — |
| 10a-IP | 1029 | 31a-IP-1 | 51a-IP | 80 | 320 | 0 | — |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 901 | 902 | 903 | 904 | 905 | 906 | 907 | 908 |

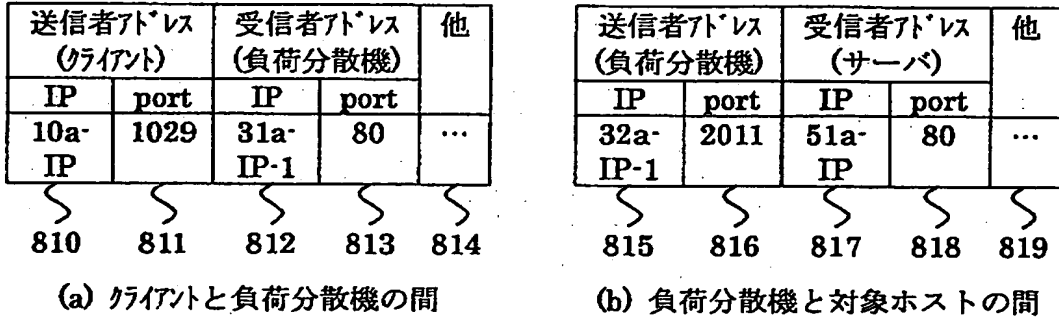
【図 1 0】

図 1 0

| クライアント | | 負荷分散機 (外部) | サーバ | | TCP フロー制御 | | |
|--------|------|---------------|--------|------|-----------|-------|-----------|
| IP | port | IP | IP | port | seq. | delta | timestamp |
| 10b-IP | 1028 | 31a-IP-2 | 51b-IP | 21 | 29 | 2 | — |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 901 | 902 | 903 | 904 | 905 | 906 | 907 | 908 |

【図 1 1】

図 1 1

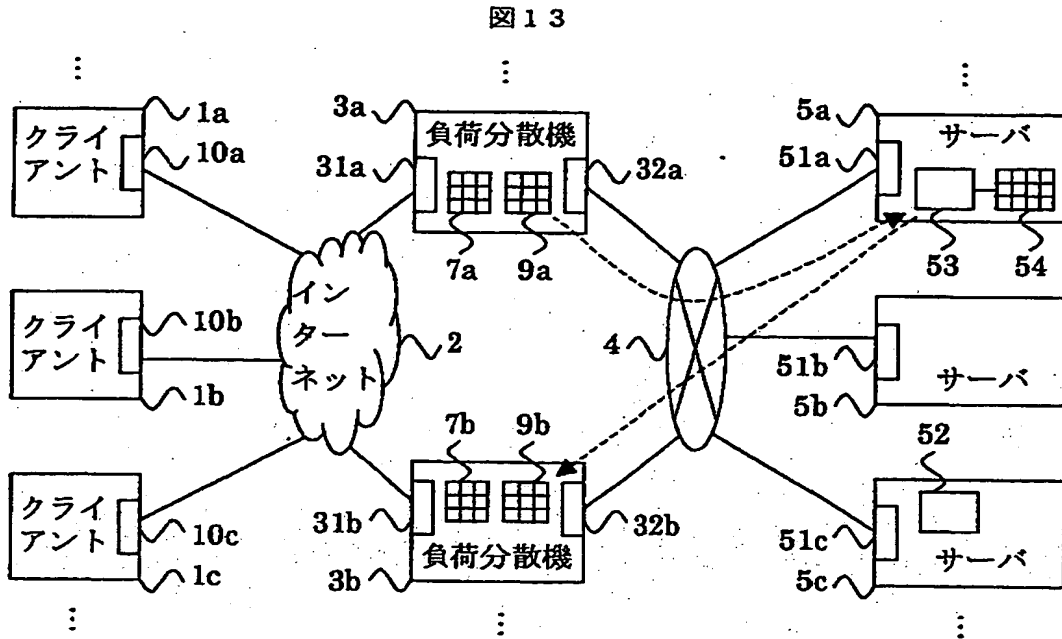


【図 1 2】

図 1 2

| クライアント | | 負荷分散機 | | | サーバ | | TCP フロー制御 | | |
|------------|------|--------------|--------------|------|--------|------|-----------|-------|----------------|
| | | 外部 | 内部 | | | | | | |
| IP | port | IP | IP | port | IP | port | seq. | delta | time- stamp |
| 10a- IP | 1029 | 31a- IP-1 | 32a- IP-1 | 2011 | 51a-IP | 80 | 320 | 0 | — |
| 10b- IP | 1028 | 31a- IP-2 | 32a- IP-2 | 2010 | 51b-IP | 21 | 29 | 2 | — |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 910 | 911 | 912 | 913 | 914 | 915 | 916 | 917 | 918 | 919 |

【図 13】



【図 14】

図 14

| クライアント | | 対象ホスト | | TCP フロー制御 | |
|--------|------|--------|------|-----------|-----------|
| IP | port | IP | port | seq. | timestamp |
| 10a-IP | 1029 | 51a-IP | 80 | 320 | — |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

930
931
932
933
934
935

【書類名】 要約書

【要約】

【課題】 複数の負荷分散機間で動的通信負荷分散もしくはフェールオーバーを実現する。

【解決手段】 クライアント1a～1cとサーバ5a～5cとのアクセスの通信と、負荷分散機3a、3b内に記録するアクセス対応表9a、9bの情報を負荷分散機の間で移動する。アクセスを負荷分散機3a→3bに移動するとき4つのステップを実行する：（1）負荷分散機3bに移動を知らせ、エラー処理を抑える。（2）通信を負荷分散機3bに移動する。（3）アクセス対応表9aの情報を負荷分散機3bに記録する。（4）負荷分散機3bを通常稼動状態に戻す。フェールオーバーを実現するために負荷分散機3aはアクセス対応表9aの情報をサーバ5aに複写する。

【選択図】 図13

認定・付加情報

| | |
|---------|---------------|
| 特許出願の番号 | 特願2001-361913 |
| 受付番号 | 50101742247 |
| 書類名 | 特許願 |
| 担当官 | 第八担当上席 0097 |
| 作成日 | 平成13年11月29日 |

<認定情報・付加情報>

| | |
|-------|-------------|
| 【提出日】 | 平成13年11月28日 |
|-------|-------------|

出 願 人 履 歴 情 報

識別番号

[000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所